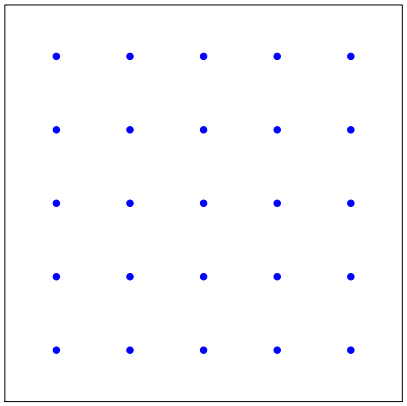
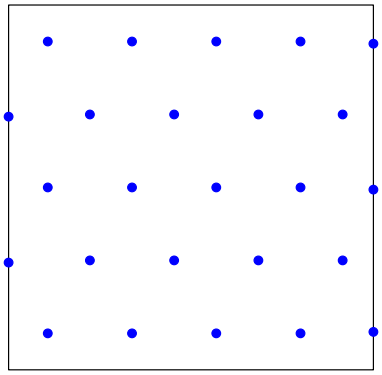
Spatial Sampling Designs

- If you can choose locations to sample, what locations?
- 2 key questions
  - 1) Single sample or sequential / multi-phase sample
  - 2) What is the study goal?
    a) good predictions across an area? Or,
    b) good estimate of variogram
- Study goals are antagonistic. Intuitively:
  - for a, want to spread out samples
  - for b, need some closely spaced values
- Predictions across the study area:
  - Good prediction when a prediction point to an obs. location.
  - So try to minimize the maximum of those distances
  - When sample on a grid
    - square, hexagonal, or herringbone (examples below)
    - Hexagonal often called triangular
    - Hexagonal has smallest maximum distance from a grid point

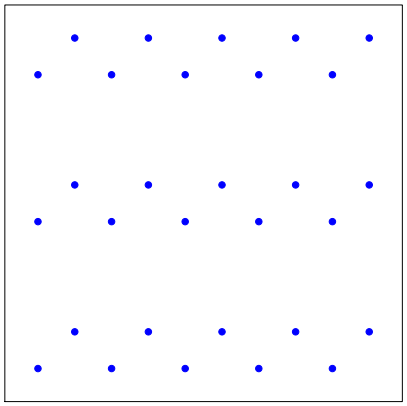
Square grid



Hexagonal grid

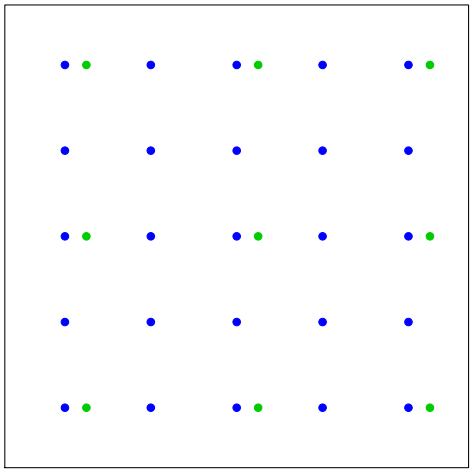


Herringbone grid



Sampling Designs

- Good estimate of SV
  - Want lots of different distances
  - especially lots of short distances, each with many pairs
    - information about semivariance at short lags
    - good extrapolation to nugget
  - square grid: many pairs separated by grid spacing, none shorter
  - consider herringbone
  - Or, add additional points to a square grid (picture next)
- How many points and how closely spaced?
  - R. Webster has done a lot of work on this
  - his examples are all soil, but the principles are general



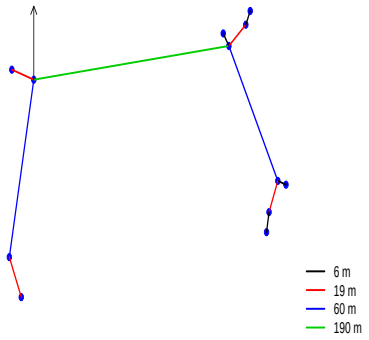
Three aspects of a sampling scheme

- 1) How many points? Webster's advice:
  - 100 obs "may be acceptable" to estimate a SV
  - at least 144 obs "seems necessary"
  - 400 obs $\rightarrow$ "great precision", "seems extravagent"
  - Other work: REML estimates require fewer obs (e.g., 50 not 100)
    - But more recent work by Webster suggests 50 definitely not sufficient
- 2) Minimum distance between pairs of obs.
  - If too large, miss short-range spatial correlation
  - Example: practical range of the SV is 50m
    - grid spacing = 33m (2/3'rds the range): SV looks like pure nugget
    - grid = 20m to 25m (ca 1/2 the range), adequate when no nugget
    - need pairs $<$ 20m to reliably estimate a non-zero nugget
- 3) Spread points out across the area to predict well
- You see the conflict between goals

Webster's nested sampling design

- Background:
  - Pre-existing survey of soils in Wyre Forest, England
  - SV was pure nugget beyond 197m
  - No pairs of obs. closer than 197m
- Want a new set of samples that provides information about short range spatial correlation and can be used to map soil properties across 26 km^2
- Unbalanced scheme designed to collect information on many different distances and directions
- Result: efficient estimation of the SV
- Sampling centers separated by 600m
- Picture of one "center"

Webster's nested sampling design



Adaptive sampling

- Concept:
  - Collect preliminary data
  - analyze sample then plan new observations
- Adaptive sampling:
  - krige surface,
  - locate places with highest prediction variance,
  - place new samples there.

Practical advice: sampling to estimate variogram

Based on Chapter 5 in Oliver and Webster, 2015
All this advice assumes isotropy

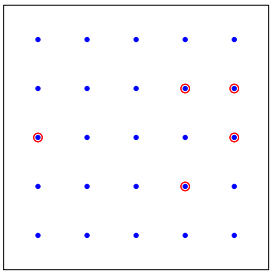
- Need a prior estimate of the correlation range (= practical range)
  - How far do you have to move to get essentially no spatial correlation?
- 1) sample so that the largest lag distance $>$ correlation range
  - Remember, variograms usually only computed to 1/2 smallest plot dimension
- 2) Want about 6 SV estimates within correlation range and 4 larger
  - That's so you get a reasonable idea of SV shape.
  - When 2 points within correl. range, any model w/nugget will fit
  - Have 2 data values and model has 2 parameters (range and nugget)
- 3) Need 100-150 locations to get a reasonably precise empirical variogram

Practical advice: sampling to map the response

- Goal: minimize the max prediction variance over the prediction grid
- equiv. to minimizing max distance between a sampling point and any other
- Smallest when sample on a grid.
  - Smallest max distance for hexagonal = triangular grid
  - Square grid almost as good, often much more practical
    - Max distance 14% larger
    - But square grid has 4 neighbors; hexagonal grid only has 3
    - Combination $\Rightarrow$ Max prediction variance for square grid only slightly $>$ hex grid
- gstat ossfim() function will calculate max pred. var. given a SV model, grid spacing, and block kriging
- Short range correlation $\Rightarrow$ closely spaced grid and lots of points
  - But at least you will be aware of the heterogeneity
- If you don't have a prior estimate of correlation range?
  - Can you use ancillary information (e.g. remote sensed spectral band)?
  - If not, everything is a guess (and no better)

Cokriging

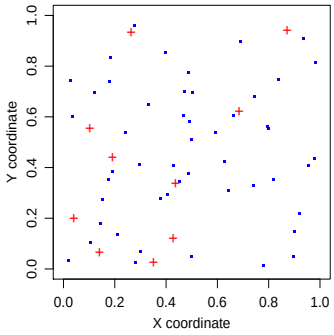
- Two (or more) spatial variables, Z_1 and Z_2
- More (usually many more) observations for Z_1 than Z_2
- Two possible relationships:
  - Z_1 and Z_2 measured at same locations, Z_1 measured at prediction points



  - Use UK with Z_1 as the covariate

Cokriging

- Z_1 and Z_2 measured at different locations



- Cokriging uses information from Z_1 to predict Z_2

Cokriging

- Concepts:
  - Each is spatially correlated: Cov $Z_1(\boldsymbol{s})$, $Z_1(\boldsymbol{s}+\boldsymbol{h})$
  - If measured at same location, are correlated: Cov $Z_1(\boldsymbol{s})$, $Z_2(\boldsymbol{s})$
  - Result: spatial cross-correlation: Cov $Z_1(\boldsymbol{s})$, $Z_2(\boldsymbol{s}+\boldsymbol{h})$
  - estimate that cross correlation by cross semivariogram:

$$\gamma_{ij}(h) = \mathrm{E}\ [Z_i(s) - Z_i(s+h)]\,[Z_j(s) - Z_j(s+h)]$$

  - i.e., how the covariance between Z_i and Z_j changes with distance
  - use to improve predictions of Z_2 and/or Z_1
- various applications
  - most promising is to use fine grid to predict a sparsely observed, but correlated, value
- Thoughts based on my limited experience:
  - needs good correlations between Z_1 and Z_2 to substantially improve predictions

co-kriging

- How actually implemented:
  - stack $Z_j(s)$ values below $Z_i(s)$ values to make one long vector
  - that stacked vector has a partitioned VC matrix:

$$\begin{bmatrix} \text{VC for } Z_i & \vdots & \text{Cov } Z_i, Z_j \\ \gamma_i(h) & \vdots & \gamma_{ij}(h) \\ \cdots & \cdots & \cdots \\ \text{Cov } Z_j, Z_i & \vdots & \text{VC for } Z_j \\ \gamma_{ji}(h) & \vdots & \gamma_j(h) \end{bmatrix}$$

  - Must satisfy some requirements (e.g. positive definite, next slide)

postive-definite matrices

- A fundamental, important property of a Variance-Covariance matrix
- Technical definition: V is a symmetric matrix
  - V is positive definite iff $a'Va > 0$ for every choice of the vector a.
- Why important:
  - V is the variance-covariance matrix of random vector Z
  - To be a valid VC matrix:
    - Var $Z_i > 0$ for each element of Z, so all elements of diag$(V) > 0$
    - AND Var of any linear combination of Z_i, e.g. $Z_1 + Z_2 - 0.5Z_3$, must be > 0.
    - that linear combination is aZ and has variance $a'Va$
  - So if V is not positive definite, it is not a valid VC matrix.

postive-definite matrices

- Example:

$$V = \begin{bmatrix} 1 & 1.5 \\ 1.5 & 1.5 \end{bmatrix}$$

  - Var $Z_1 = 1$, Var $Z_2 = 1.5$, Var $Z_1 + Z_2 = 1 + 1.5 + 2*1.5 = 5.5$
  - But, Var $Z_1 - Z_2 = 1 + 1.5 - 2*1.5 = -0.5$. OOPS!
  - V is not a valid VC matrix
- Diagnosis:
  - Calculate correlation matrix corresponding to V. All off-diagonal elements between -1 and 1
  - Better: calculate eigenvalues of V. All are > 0.
  - Some eigenvalues $= 0$ means that some correlations $=$ -1 or 1, linear dependencies among variables.

Linear model of coregionalization

- So need VC matrix for Z to be positive definite
  - Hard to do for arbitrary $\gamma_i(h)$, $\gamma_j(h)$, and $\gamma_{ij}(h)$
  - Can force V to be positive definite by using linear model of coregionalization
    the three SV's have same nugget, same range, positive definite matrix of sill values
- co-kriging uses cross-variogram to make predictions
- same goal minimize MSEP, same form of predictor